

CLONAREA INFORMAȚIONALĂ

Ion IVAN
Paul POCATILU
Marius POPA
Mihai SACALĂ

Academia de Științe Economice, București

Rezumat: Se definește conceptul de clonare informațională pentru a stabili măsura în care diferite componente hardware și/sau software sunt reutilizate în alte sisteme de calcul și/sau sisteme de aplicații. Sunt prezentate diverse metode de identificare a clonelor în componentele informatice.

Cuvinte cheie: clone informatice, componente informatice clonate.

1. Componente informatice

Informatica modernă presupune integrarea componentelor hardware și software într-o infrastructură cu grad de complexitate ridicat. Problematika reproductibilității, a posibilităților de reutilizare are particularități determinate de efortul generat de multiplicarea exemplarelor.

Componentele informatice sunt de naturi și utilizări diferite.

Documentațiile sunt produse destinate prezentării de componente hardware, de programe și de servicii. Documentațiile însoțesc produsul sau serviciul și sunt structurate încât să prezinte:

- descrierea pe scurt a produsului sau serviciului;
- licența produsului;
- necesarul minim de resurse;
- mod de instalare;
- funcțiile;
- mod de deinstalare;
- modul de achiziționare al produsului sau serviciului – informații despre producător, preț – în cazul produselor distribuite ca shareware sau freeware;
- informații privind asistența tehnică, acordată de producător.

Componentele hardware sunt destinate realizării unor funcții de memorare pe disc magnetic, disc optic, memorie electronică, bandă magnetică, imprimantă, hard-disk, de preluare informații prin scanare, de transmitere date în rețea și de dirijare spre noduri prin hub-uri.

Banda magnetică – biții de informație sunt memorați prin variațiile câmpurilor magnetice.

Discul magnetic – biții de informație sunt memorați, de asemenea, prin variațiile câmpurilor magnetice de pe disc. Discul este fix (*hard*) sau flexibil (*floppy*).

Discul optic – biții sunt memorați ca pit-uri. Citirea se realizează cu laser. Această informație este fie *read-only*, *write-once read many*, fie reprogramabilă.

Memorie electronică – biții sunt memorați cu ajutorul unor circuite electronice, prin impulsuri electrice. Această memorie se folosește pentru stocarea datelor în memoria internă a calculatorului.

Calculatoarele comunică într-o rețea locală LAN (Local Area Network) sau rețea globală WAN (Wide Area Network).

O rețea LAN reprezintă o colecție de calculatoare conectate într-o arie restrânsă, prin conexiuni electronice comune (*backbones*). O rețea LAN permite:

- partajarea resurselor, în primul rând a datelor;
- administrarea ușoară a utilizatorilor.

Rețelele LAN sunt interconectate direct sau prin intermediul unei rețele WAN.

O rețea WAN conectează rețele pe o suprafață mare între calculatoarele aflate în diferite clădiri, orașe sau țări.

Hub – punct de conexiune comună pentru diverse dispozitive într-o rețea. Este folosit pentru a conecta segmente de rețele de tip LAN, conținând mai multe porturi. Când datele sosesc la un port sunt copiate la celelalte porturi astfel încât toate segmentele rețelei pot vedea datele.

Software-ul de aplicație este destinat soluționării claselor de probleme de optimizare, de rezolvare a sistemelor liniare, de estimare a coeficienților modelelor econometrice, de calcul al indicatorilor statistici.

Bazele de date reprezintă seturi sau colecții de date organizate pe niveluri, în mod coerent, asigurând integritatea și securitatea, structurate conform modelelor de date construite, cu o redundanță minimă și controlată prin tehnici de proiectare, accesibil mai multor utilizatori în timp util [8], [3]. Componentele bazei de date sunt:

- colecțiile de date propriu-zise;
- dicționarul de date;
- fișierele auxiliare;
- datele virtuale (viziuni, zonele de memorie).

Tipuri de modele de date fundamentale pentru baze de date:

- arborescent;
- rețea;
- relațional;
- orientat pe obiecte.

Fișierele sunt colecții de date sau programe organizate după criterii calitative, de prelucrare și de scop.

Din punctul de vedere al unui program, un fișier este o structură de date externă, formată din elemente numite înregistrări. Dimensiunea fișierelor este limitată numai de dimensiunea fizică a suportului utilizat. Pe un disc flexibil sau pe un disc fix sunt memorate mai multe fișiere.

Utilitarele sunt produse software, realizate cu scopul de a îmbunătăți gradul de folosire a resurselor unui sistem de calcul.

Compilatoarele sunt instrumente software, care conduc la obținerea modulelor obiect, necesare pentru realizarea prin editare de legături a programului executabil. Intrările compilatoarelor sunt fișiere sursă scrise într-un limbaj evoluat de programare.

Sistemele informatice reprezintă un ansamblu de dispozitive hardware și de produse software, interconectate, care funcționează în vederea satisfacerii anumitor cerințe ale utilizatorilor.

Serviciile în sens larg, sunt fapte, acțiuni care avantajează, satisfac exigențele celor care le solicită și sunt și beneficiari direcți. În cazul de față, serviciul reprezintă o acțiune efectuată cu ajutorul unui produs software la cererea unui utilizator, în situația în care sunt folosite rețelele de calculatoare. Acțiunea constă în furnizarea de informații, efectuarea de operații de înregistrare, interogare, căutare, ștergere, transfer, alocare de resurse și efectuare de plăți.

2. Funcția de copiere

Cu ajutorul instrumentelor de gestiune a fișierelor, se realizează următoarele operații: listarea fișierelor de pe disc, aranjarea fișierelor în directoare și ștergera fișierelor. În plus, se aplică operații de formatare, copiere și etichetare a discurilor, de sortare și de modificare a fișierelor.

Sistemele de operare, oricare ar fi acestea, sunt înzestrate cu funcții de manipulare a fișierelor precum:

- *creare fișier* – operație de alocare a unei zone de memorie pe suportul de memorie extern unde se introduc și se stochează datele fișierului;
- *modificare fișier* – constă în schimbarea unuia sau a mai multor câmpuri din fișier; se realizează la nivelul numelui de fișier sau de structură; modificarea la nivelul numelui de fișier se numește *redenumire* de fișier;
- *sortare fișiere* – constă în așezarea înregistrărilor într-o anumită ordine definită aprioric în funcție de câmpurile fișierului și importanța acestora.

Sistemul de operare MS-DOS permite copierea fișierelor prin instrucțiunea *copy* definită astfel:

copy [/a | /b] sursă [/a | /b] [[/a | /b] + sursă [/a | /b] [+ ...]] [/v] [/n] [/y | /-y] [/z] [/a | /b] [destinație [/a | /b]]

unde parametrii au semnificația următoare:

sursă – specifică locația și numele unuia sau mai multor fișiere de unde se copiază; locația este un disc, un director, un fișier sau o combinație de acestea;

destinație – specifică locația și numele unuia sau mai multor fișiere unde se copiază; locația este un disc, un director, un fișier sau o combinație de acestea;

/a – indică un fișier text ASCII;

/b – indică un fișier binar;

/v - verifică dacă noile fișiere sunt scrise corect;

/n – se folosește un nume scurt de fișier când copierea are loc pentru un fișier cu un nume mai mare de opt caractere sau extensia e mai mare de trei caractere;

/y – se cere confirmarea suprascrierii unui fișier;

/z – se copiază fișiere din rețea în mod restartabil, dacă în timpul copierii se pierde conexiunea.

Sistemul de operare Windows efectuează copierea de fișiere prin următoarele metode:

- a) "trageră și plasare" (drag-and-drop) – sunt efectuate următoarele operații:
- se selectează fișierul care se copiază;
 - se ține apăsat butonul stânga al mouse-ului;
 - se "plasează" fișierul deasupra directorului destinație;
 - se eliberează butonul stânga al mouse-ului.
- b) "copiere și lipire" – sunt parcurse următoarele operații:
- se selectează fișierul care se copiază;
 - se copiază fișierul în *Clipboard* prin selectare opțiune *Copy* din meniul de context, meniul Edit sau bara de instrumente;
 - se selectează destinația unde se copiază fișierul;
 - se copiază propriu-zis fișierul prin selectare opțiune *Paste* din meniul de context, meniul Edit sau bara de instrumente.
- c) "trimitere la disc flexibil" - dacă se copiază fișiere pe dischete, atunci, din meniul de context, se alege opțiunea *Send To 3 1/2 Floppy (A)*; trimiterea este generalizabilă.

Fișierele se copiază fie unul câte unul, caz în care sunt repetate operațiile de mai sus pentru fiecare, fie toate o dată, acestea selectându-se cu ajutorul mouse-ului sau a tastelor de directare, pe de o parte, și a tastelor CTRL sau SHIFT, pe de altă parte.

De exemplu, dacă fișierul A este copiat, după efectuarea operațiilor descrise anterior, se obține fișierul B. Copierea se realizează:

- pe același suport;
- pe alt suport;
- în același director;
- în alt director.

Copierea fișierelor reprezintă unul din cazurile de clonare perfectă.

- *tipărire fișiere* – presupune realizarea imprimării conținutului fișierelor text cu sau fără format;
- *ștergere fișiere* - operație prin care se realizează eliminarea de pe suportul magnetic extern a conținutului și a datelor de descriere.

Problema gestionării documentelor sub aspectul denumirii, versiunii, lungimii, datei creării și/sau modificării de fișiere constituie un domeniu distinct deosebit de important, abordat de cercetările în birotica modernă.

Există instrumente care permit clonarea dischetelor, a discurilor optice și a hard-disk-urilor. Pe lângă faptul că se copiază fișierele și directoarele de pe o unitate de stocare pe alta, având același conținut, clona obținută are aceeași structură ca și originalul.

3. Elemente informatice

Fiecare componentă informatică se descompune în părți care sunt manipulate funcție de obiectivele urmărite.

Astfel, pentru *fișiere* se selectează o parte din fișier și se copiază în alt fișier.

De exemplu, se consideră fișierul F_1 care are părțile A, B, C și D. Se selectează partea B și se copiază în fișierul F_2 care conține deja părțile X, Y și Z. Fișierul F_2 modificat conține X, Y, Z, B.

În situația în care în fișierele F_1 și F_2 se găsesc părți care intersectate sunt diferite de mulțimea vidă, rezultă că aceste fișiere au componente identice, respective clone.

Clonele se obțin prin copiere sau macroexpandare.

Se consideră un *program* care apelează procedurile P_1 și P_2 . Dacă programul considerat apelează procedurile P_1 și P_2 , dar acestea nu sunt incluse în bibliotecă, ci textele lor sunt inserate în fișierul sursă, procedurile P_1 și P_2 sunt clone.

În cazul în care programul folosește macrodefiniția:

```
add4 macro a, b, c, d, suma
```

```
    push ax
    mov ax, a
    add ax, b
    add ax, c
    add ax, d
    mov suma, ax
    push ax
```

```
endm
```

această macrodefiniție fiind apelată în programul considerat:

```
...
add4 x, y, z, w, T1
```

```
...
```

și în alt program avem:

```
...
add4 m, n, o, p, T2
```

```
...
```

prin expandare s-au obținut clone, adică secvențe identice ca operatori, dar diferite ca operanzi.

Bibliotecile de funcții, de clase stau la baza reutilizării software-ului, iar clonarea presupune includerea de secvențe din bibliotecă prin copiere și adaptare ca părți ale produsului-program, spre deosebire de referirea de componente din bibliotecă în mod natural, în concordanță cu facilitățile limbajelor de programare. Dacă într-o bibliotecă există funcția *tip* $f(p_1, p_2, \dots, p_n)$ referirea $x = f(a_1, a_2, \dots, a_n)$ nu reprezintă clonare. În schimb, dacă funcția conține secvența de instrucțiuni:

```
if (a == b) x = 3;
    else
    x = 4;
```

Și în programul elaborat se preia această secvență exact așa cum este în funcție sau înlocuind variabilele a, b, x, cu alte variabile, dar menținând prelucrările, procedul reprezintă clonare a secvențelor de program.

În *baze de date* problema clonării se pune diferențiat:

- clonare structură și inițializare cu conținut diferit;
- clonare conținut.

Clonarea structurii unei baze de date constă în folosirea în altă bază de date a acelorași câmpuri având sau nu aceeași denumire, dar descrierea câmpurilor fiind identică.

De exemplu, se consideră o tabelă T_1 dintr-o bază de date cu următoarea structură:

Tabelul 1

Denumire câmp	Tip	Dimensiune
marca	number	
nume	char	15
prenume	char	25
varsta	number	

Dacă se construiește o tabelă T_2 cu structura:

Tabelul 2

Denumire câmp	Tip	Dimensiune
marca_ang	number	
nume_ang	char	15
prenume_ang	char	25
varsta_ang	number	

atunci se spune că tabela T_2 este o clonă a tablei T_1 .

Clonarea conținutului unei baze de date presupune copierea de n-tupluri (valori ale câmpurilor definite) într-o altă bază de date.

De exemplu se consideră următoarele înregistrări în tabela T_1 :

Tabelul 3

Marca	Nume	Prenume	Varsta
10	Ionescu	Ion	35
24	Popescu	Gheorghe	42
13	Vasilescu	Ioan	40

și în tabela T_2 :

Tabelul 4

Marca_ang	Nume_ang	Prenume_ang	Varsta_ang
10	Ionescu	Ion	35
31	Constantinescu	Gheorghe	37

Se remarcă faptul că, în tabela T_2 , există un tuplu care are aceleași valori ca în tabela T_1 , fiind, deci, o clonă.

Clonarea fișierelor este un caz particular de clonare a bazelor de date. Este important ca, în fișierele considerate originale, să fie identificate elemente caracteristice precum: număr particular de repetări a delimitatorilor, erori întâmplătoare sau voite, omisiuni în completarea unor câmpuri etc. Evidențierea clonării se stabilește indubitabil dacă un alt fișier conține exact aceleași elemente caracteristice, pe care le are fișierul original numit și fișier de raportare.

Problema clonării constă în identificarea de părți comune în mai multe *documentații*, fragmentul respectiv fiind o clonă. De fapt, în acest caz, clonarea intervine atunci când produsul software, descris prin documentație, este obținut în urma procesului de clonare.

În cazul *serviciilor*, clonarea constă în multiplicarea acțiunilor de furnizare de informații, de înregistrare și interogare în baze de date, de căutare și de transfer a informațiilor, efectuate pe mai multe noduri din rețea.

4. Evidențierea clonelor informatice

Indiferent de forma de existență a produselor și serviciilor în domeniul informaticii, evidențierea și gestionarea procesului de clonare reprezintă o activitate importantă întrucât limitarea clonelor creează premise obținerii de software performant și de atragere de resurse financiare pentru dezvoltarea producției și calificarea personalului. În acest scop, se impune dezvoltarea de instrumente software, care să analizeze arhitecturi hardware, produse software și baze de date existente într-o arie de utilizare precis delimitată și să identifice clonele existente. În cazul în care clonele și procesul de clonare depășesc limitele definite printr-un cadru legal existent la un moment dat, se procedează la efectuarea de corecții pe toate planurile.

Astfel de produse software sunt realizate deja pentru stabilirea părților comune din texte redactate cu procesoarele de tip Word și pentru identificarea clonelor din programe C++ precum și depistarea elementelor de proporționalitate, care generează clone în tabele construite prin prelucrări de tip matriceal [1].

Absența clonelor este specifică datelor ortogonale [2]. Produsul *ident_soft* este destinat evaluării gradului de ortogonalitate a fișierelor text. Algoritmul constă în următorii pași:

- se traversează fișierul F_i și se construiește vocabularul și frecvențele de apariție;
- se traversează fișierul F_j și se construiește vocabularul și frecvențele de apariție;

- se stabilește vocabularul comun;
- se aleg frecvențele minime;
- se calculează indicatorul l ca raport între suma de frecvențe minime și suma de frecvențe maxime.

Expresia analitică a indicatorului l este:

$$l = \frac{\sum_{i=1}^n f_{\min_i}}{\sum_{i=1}^n f_{\max_i}}$$

unde:

n – numărul de cuvinte care formează fișierele;

f_{\min_i} – frecvența de apariție minimă a cuvântului i care apare în fișiere;

f_{\max_i} – frecvența de apariție maximă a cuvântului i care apare în fișiere.

În cazul organizării de competiții problemele de ortogonalitate sunt esențiale.

5. Concluzii

Clonarea este necesară până la un punct, dar este periculoasă pentru că nu mai generează dezvoltarea de software nou.

Prin clonarea neautorizată nu se acoperă cheltuielile de către firmele care au produs software. Produsele software clonate neautorizat apar pe piață cu prețuri foarte mici.

Bibliografie

1. **IVAN, I., M. POPA, M. SACALĂ:** Ortogonalitatea datelor. În: Revista Română de Statistică (în curs de apariție).
2. **IVAN, I., P. POCATILU, M. POPA:** Analiza cantitativă în managementul proiectelor. În: Studii și Cercetări de Calcul Economic și Cibernetică Economică (în curs de apariție).
3. **LUNGU, I., M. VELICANU, M. MUNTEAN:** Teste și probleme de programare în Visual FoxPro și Oracle, Editura Petron, București, 2000.
4. **MATTHEWS, M., B. DOBSON:** MS-DOS și Windows, Editura Teora, București, 1996.
5. **SMEUREANU, I., I. IVAN, M. DÂRDALĂ:** Structuri și obiecte în C++, Editura CISON, București, 1998.
6. **STANCIU, E.:** Reutilizarea componentelor program în sisteme complexe (referat doctorat), București, 2000.
7. **STANCIU, E.:** Căi de creștere a gradului de reutilizare a componentelor software. (referat doctorat), București, 2000.
8. **VELICANU, M., I. LUNGU, C. BODEA, C. IONIȚĂ, G. BĂDESCU:** Sisteme de gestiune a bazelor de date, Editura Petron, București, 2000.